Evaluation of terrestrial LiDAR technology for the development of local tree volume equations

Demetrios Gatziolis[a], Sorin Popescu[b], Ryan Sheridan[b], and Nian-Wei Ku[b]

[a] USDA Forest Service, PNW Research Station, Resource Monitoring & Assessment, Portland, Oregon, dgatziolis@fs.fed.us
[b] Department of Ecosystem Science and Management, Texas A&M University

Abstract

Tree volume, a key parameter in all forest inventory systems, is typically calculated by using species-specific polynomial equations which are constructed by using measurements obtained from destructive tree sampling. Owing to cost, logistic, and sometimes policy constraints, typically only a small number of trees within the often large, multiregional distribution range of a species is destructively sampled. In areas characterized by diverse ecological, climatic, and forest structure conditions, such as the US Pacific Northwest region, this practice has lead to the development and use of volume equations whose prediction accuracy and precision at the local level is often unknown, even for common tree species. The detailed representations of tree stems offered by modern terrestrial LiDAR technologies are sometimes assumed to be an alternative, economical, non-destructive source of stem dimensionality information capable of supporting the development of local tree volume equations. We evaluate this assumption by comparing the merchantable volume of trees calculated from multiple diameter measurements obtained by experienced tree climbers at short intervals along the stem of standing trees in eastern Oregon to the volume estimated by processing corresponding terrestrial LiDAR data. Our methodology attempts to reconstruct the tree stem by exploring the mathematical morphology of terrestrial laser point clouds in voxel space and by utilizing elements of the graph theory. We found that if at least one stem segment near the upper end of the tree and a few other segments below are not obstructed when viewed from the laser scanning station, volume estimation errors no larger than 2% can be expected. For most trees, however, poor representation in the point cloud of the upper parts of the stem leads to an often substantial underestimation of volume. As expected, the accuracy of volume estimates improves with higher point densities.

Keywords: Terrestrial LiDAR, tree volume, voxel morphology

1. Introduction

Estimation of individual tree volume is a key component of forest inventory systems. Volume estimates are obtained via species- or species-group-specific equations that usually comprise two independent variables, tree diameter at breast height (DBH) and height, both characteristic, easy-to-measure descriptors of tree size. Volume equations come in a variety of forms including polynomial curves that approximate the tree diameter at regular intervals along the tree stem, or represent integrations of taper functions (Max & Burkhart, 1976). Pending on the specific portion of the woody tissue of a tree targeted, these equations can be used to predict the merchantable volume to a minimum stem diameter, the total volume of the main stem (rarely), or the volume with or without bark. Volume equations often form the basis for predicting tree biomass by using known specific gravity values for individual tree species.

The traditional approach used for developing volume equations involves destructive sampling. A set of trees believed to represent the targeted tree-species population is identified and the trees are subsequently felled. Diameter measurements made at regular, short intervals along the felled stems provide the data for developing taper functions and ultimately volume equations. The logistical complexity and cost associated with the traditional approach often hinders efforts aiming at the development of new or the refinement of existing volume equations fit to local conditions. The next biggest impediment in the development of locally-unbiased tree volume equations in many regions is regulations that prohibit the

harvesting of trees, or at least of trees over a certain size, and, hence, of obtaining a representative sample of the local tree population for a given species. In such circumstances, non-destructive-tree-sampling approaches could a useful, and likely the only viable, alternative for forest mensuration purposes, assuming of course that they meet accuracy and precision standards.

One such alternative comes in the form of terrestrial light detection and ranging (TLiDAR), a technology capable of generating abstract representations of objects illuminated with laser pulses from one or more near-ground locations. It has been increasingly used lately for forest applications including estimation of dendrometric parameters (Hopkinson *et al.*, 2004) and the assessment of tree-stem dimensionality (Thies *et al.*, 2004). TLiDAR data comprise a usually dense cloud of pulse returns or points from object surfaces precisely georeferenced in three dimensions. Processing of the point cloud could yield detailed representations of tree elements that range in size from the main stem to small twigs or leaves.

Pfeifer *et al.* (2004) managed to derive the structural architecture of trees scanned from multiple stations in the form of skeletons constructed by exploring the three-dimensional mathematical morphology in voxel space and they estimated stem thickness by fitting cylinders to skeleton sections. Many tree branches, especially those in the upper parts of the crowns were not reconstructed. Tree skeletons can also be derived by first organizing the point cloud into a geodesic graph (Verroust and Lazarus, 2000) which constitutes a three-dimensional network of points, each connected to a pre-selected number of closest neighbors. Given a reference point within the cloud, the minimum-distance path that connects each point to the reference can be computed. Connecting the centroids of points that belong to the same minimum-distance class produces the skeleton (Xu *et al.*, 2007). A variant of this approach designed to support automation was implemented by Cheng *et al.* (2007) who applied it to laser points pertaining to two small deciduous trees at leaf-off conditions collected from a single scanning station. For one of the two trees manual editing was necessary to obtain an accurate reconstruction of branch structure. Bienert *et al.* (2007) used diameter profile fitting on tree stems the locus of which had been previously identified via segmentation of point density rasters and then applied a set of heuristics labeled 'reliability factors' to detect and correct diameter over- or underestimation. More recently Lefsky and McHale (2008) estimated the stem, branch volume and canopy volume of 179 deciduous trees with complex structure scanned at leaf-off conditions in an urban setting. By examining the connectivity of voxels that contained at least one point, they first identified clusters of connected voxels. Processing of the clusters, application of heuristic rules, and manual editing yielded representations of the main stem and for the majority of larger branches, and volume was computed by cylinder fitting. The authors noted that computer memory limitations often enforced voxel resolution coarser than what it was considered optimal and that processing time requirements constrained recursive application of heuristic rules. As a surrogate of tree volume, they compared stem diameter measurements obtained with an optical dendrometer to corresponding diameters of fitted cylinders and reported high correlation between the two (R^2 between 96 and 98 percent).

The objective of this study was to evaluate the utility of TLiDAR data for assessing stem diameter and volume of trees represented by a range of sizes and belonging to two coniferous species. The evaluation process combined elements of the approaches mentioned above and it was based on *in-situ* measurements of stem diameter at regular intervals recorded by experienced tree climbers.

2. Methods

2.1. Field data

Nine fixed-radius (7.32 m) subplots in 4 Forest Inventory and Analysis (FIA) plots located in the Malheur National Forest of eastern Oregon, USA, dominated by coniferous tree species (*Ponderosa pine* and *Abies grandis*) were visited in August of 2009. Omnidirectional laser scanning was performed by using a Leica ScanStation II instrument positioned on each subplot center. The specifications of the instrument are

reported in Table 1. Nominal angular resolution of 0.1146^{o}, or 2 cm at 10 m range, was used during the scanning. A selected subset of subplot trees, and five other trees outside the subplot limits but within the FIA hectare plot footprint (56.4 m radius) were scanned from additional stations with field of view just wide enough to include the entire tree and with finer angular pulse resolution. A total of 10 trees received high-density scans, and 6 of them were scanned from two different stations. Point data from independent scans targeting the same tree were co-registered automatically using proprietary software and reference targets placed appropriately within the scanning scenes. In a subsequent visit in early October 2009, the FIA database records from the latest regular inventory plot visit pertaining to each of the 9 subplots were updated using high precision field measurements that included DBH, height, and the location relative to the subplot center of all subplot trees. Ten of the 78 trees with DBH greater than 12.6 cm in the 9 subplots, and all other 10 trees scanned with high pulse densities were climbed by experience personnel who measured the over-bark diameter of the main tree stem at short, 1 to 1.5 m intervals. The measurements continued upwards the stem until the merchantable diameter threshold (10.2 cm) was reached. A distance tape was used to record the distance along the main stem at which each diameter measurement was obtained. Bark thickness was measured at every fifth diameter measurement. For trees growing on a slope, ground elevation differences between the base of the stem facing the scanning station and the uphill side were recorded, because the latter serves as reference for determining the elevation of the DBH measurements. All 20 climbed trees, a representative sample of the local distribution of tree sizes assessed by using data from the FIA databases, were free of structural deformities and in plain view from the scanning stations.

Table 1. Specifications of Leica ScanStation II instrument

Type	Dual axis compensated pulse laser
Pulse wavelength	532 nm (green)
frequency	< 50 kHz
footprint	6 mm at 50 m
range	300 m at 90% target albedo
Range precision	4 mm
Field of view	360^{o} horizontal, 270^{o} vertical
Scanning pattern	Top-down, bottom-up sequential
Intensity quantization	12-bit

2.2. Assessment of stem diameter and volume

The processes used to assess stem diameter and volume entailed determining the locus of individual tree stems in the point cloud, identifying the laser points belonging to each targeted tree, generating a tree representation in voxel space, processing the voxels to determine connectivity and extract the stem and branch architecture, and finally computing the dimensionality of the stems. All computations were performed using libraries ‘graph’ (Gentleman *et al*., 2010) and ‘RBGL’ (Carey *et* al., 2010) available in the R software (R Development Core Team, 2010) with external calls to custom scripts written in the C programming language.

For all omnidirectional scans a raster (25 cm cell resolution) ground surface was computed using the TiFFs software (Chen, 2007) and then used to convert point elevation to height values. Point subsets that included only returns no higher than 3 m from the ground were processed to compute point frequency (count) rasters of same resolution and co-registered to the ground surface rasters. The approximate, near-ground, two-dimensional location of tree stems was then retrieved as the center of cell clusters in the count rasters with values much higher than those of surrounding raster cells. All laser points located horizontally within 1 m from the center of each cell cluster and with height value up to 3 m were subsequently identified. These groups of points represented the portion of each tree stem just above

ground. For each point group, a single reference point in the scanner-facing direction of the stem and at DBH reference height was manually identified in point visualizations. In directional scans the reference point was selected directly from the point cloud. For each reference laser point, all other laser points horizontally within an arbitrary radius were selected, thereby forming point sets equal to the number of identified trees. The length of the radius was determined by inspecting visualizations of the corresponding point cloud, and was always large enough to include the tree stem in its entirety along with most of the branches, or all of them for free-standing trees (Figure 1a). A voxel representation of the point set was then generated for each tree, with each individual voxel labeled '1' if it contained one or more points, and '0' if empty. The resolution of the representation or, equivalently, the size of each voxel side, was determined by considering the angular resolution of the pulses and the distance between the scanner and the furthest voxel; it ranged between 8 and 15 cm. Note that excessively fine resolutions would produce a large number of filled voxels without any non-empty voxels as immediate neighbors, while with too coarse a resolution there would be too few empty voxels. Controlling the voxel size to ensure an adequate balance between filled and empty voxels was a critical prerequisite for retrieving the stem and branch architecture of the tree.

Each tree's voxel representation was processed with a variant of the connected components algorithm (Lumia *et al*., 1983). By examining the 26 immediate neighbors of each voxel, the algorithm identified clusters of filled voxels where each member voxel of a cluster had at least one filled voxel as neighbor. Once all clusters had been identified, a unique identifier for each voxel, the coordinates of its center, and the identifiers of and Euclidean distance in three dimensions to all neighboring voxels were organized in files, one for each cluster. In the standard implementation of the algorithm, computer memory availability often constraints the size of the voxel representation that can be processed. In our variant, the algorithm operates in partitions of the representation and thus it can process very large number of voxels, such as those constructed from dense scans of large trees. Owing to the explicit topology present in each voxel representation, neighbor identification and distance computations are far simpler and less demanding on computing resources compared to working directly with the laser cloud points.

The information embedded in each of the cluster files mentioned above is in essence a graph G which can be expressed as G = (V, L), where V is a vector of voxel labels or nodes, and L is a list of minimum one to a maximum of 26 links or edges to the neighbors of each element in V. The distance between two neighboring nodes V_i and V_j can be retrieved from the cluster files and attached to the corresponding link L_{ij}. Since each cluster represented by G was identified by using the connected components algorithm, any V_j can be reached by any V_i by querying G to determined the sequence of links, or path, that has to be followed. Of the usually large number of paths that link two non-neighboring elements of G, of practical interest is the one that minimizes the distance or sum of weights of the visited links. In our application we computed the minimum distance path between the reference node V_r and all other nodes (Figure 1b) using Dijkstra's (1959) algorithm, calculated the number of shortest paths that pass through each node, and then we extracted the paths from V_r to the nodes that define the convex hull of G (Figure 1c). The collection of extracted paths resembles the stem and branch architecture of the tree. Tracking the nodes where two or rarely more paths merge starting from the reference node and progressing upwards the tree while considering the number of minimum paths crossing the nodes adjacent to those examined, leads to the identification and subsequent 'pruning' of the branches and to the subset $G_s = (V_s, L_s)$ of G that contains the nodes and links that represent the main stem of the tree.

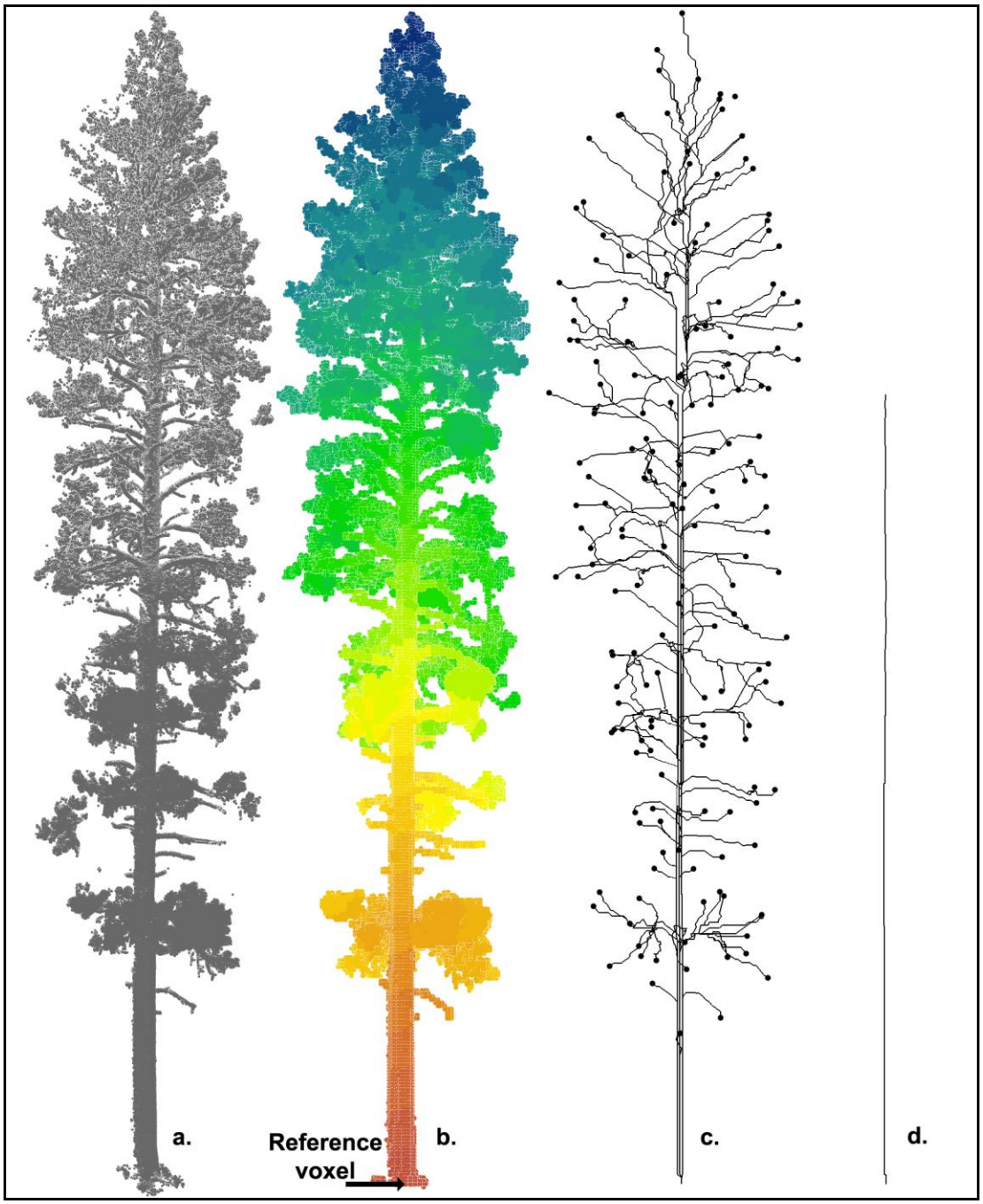

Figure 1. a. Subset of laser point cloud from dense directional scanning, b. Voxel representation colored by Dijkstra's distance from reference voxel at the base of the tree. Cooler colors indicate larger distances. c. Voxels defining the tree's convex hull denoted by dots and corresponding shortest paths to the reference voxel, and d. retrieved stem of the tree.

The sequence of nodes V_s and links L_s in G_s often exhibits localized meandering, especially in segments along the tree stem partially obstructed by branches and foliage situated between the stem and the scanning. Most of this meandering can be removed by fitting a spline through G_s (Figure 1d). In the last step of the process, all points in the immediate vicinity of any G_s element are identified and a sequence of short, often overlapping cylinders is fit to them so that the axis of each cylinder is in parallel to the

corresponding spline section. The length and width of each cylinder is evaluated repetitively to minimize the chance that its dimensionality is influenced by points that actually belong to branches or foliage. For cylinders that pass the consistency test, the diameter and the location of the middle of the axis is recorded. Finally the volume of the stem is calculated as the sum of frustum of cones formed by the diameters calculated along the stem's length.

3. Results and discussion

For 28 of 78 subplot and 5 hectare plot trees, approximately one in three, occlusion by other trees or insufficient laser point density precluded assessment of DBH. All 28 trees were scanned with standard density from the subplot center. The relationship between measured and laser-estimated DBH is shown in Figure 2. As expected, the precision of estimates obtained by using high-density scans is higher compared to those obtained by using standard density. For trees scanned with standard density, precision appears to

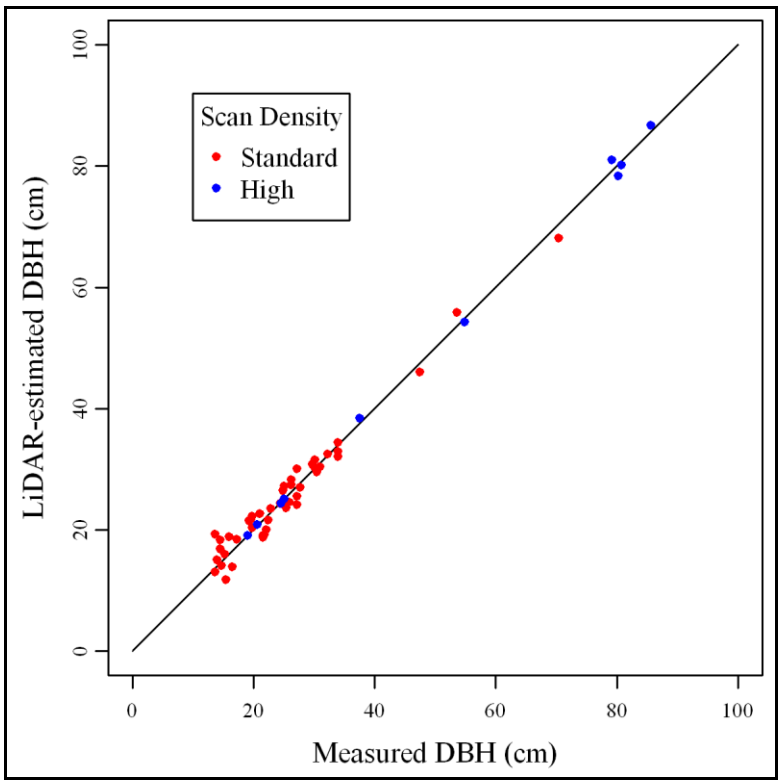

Figure 2. Relationship between measured and LiDAR estimated DBH.

decrease with diameter size. With the standard scan density used in this study, a stem at 10 m from the scanning station and with a diameter of 15 cm would be represented with a mere 7.43 points at a scan line across its axis. A planimetric assessment of diameter for this tree, similar to the one traditionally obtained by using dendrometers, would lead to a 7% underestimation of DBH. In theory, this negative bias is avoided by the cylinder fitting process, which capitalizes on the variability in the distance from the scanner to the cylinder surface to compute the radius of the cylinder and, hence, the diameter of the tree. For the tree in question, however, the expected standard deviation of distance to a horizontal cross section of its main stem at DBH elevation is only 1.65 cm which makes the computed cylinder radius particularly susceptible to noise in the measurement of range by the laser instrument or due to small objects near the surface of the stem. The effect of these error sources is less pronounced for biggest trees; the expected standard deviation of distance to the stem for a 40 cm DBH tree is 4.39 cm and for an 80 cm tree is 5.59

cm, well above the noise level one would anticipate under normal conditions. On the other hand, departure from stem roundness may introduce error in the assessment of diameter for larger trees (Figure 2).

Stem reconstruction succeeded for all 20 climbed trees, but the accuracy of the reconstruction varied. With one exception, the volume of all trees scanned with standard density was underestimated substantially (mean -21.29%, std 10.25%) (Figure 4a-c). The volume of those scanned with high density

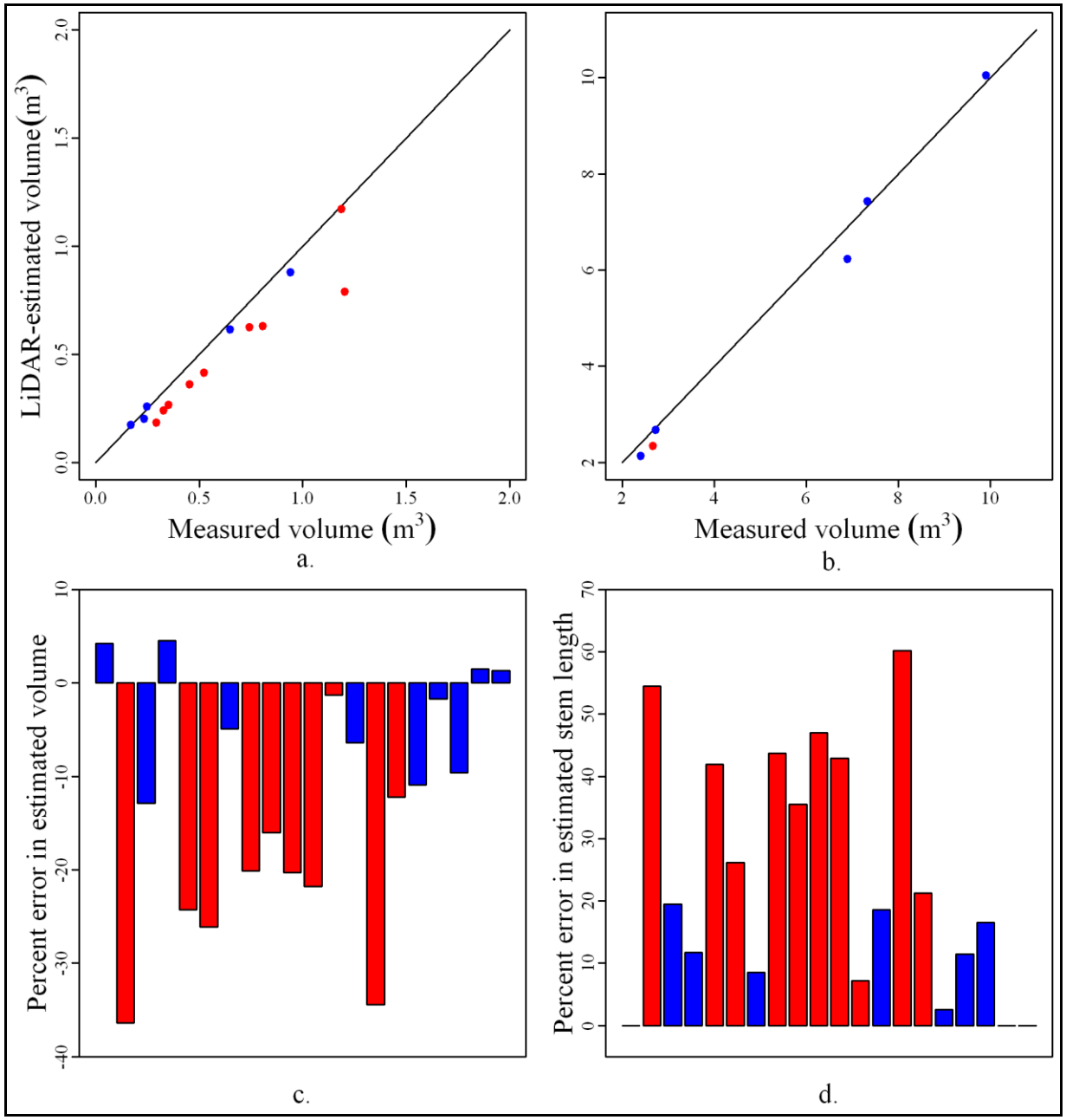

Figure 3. a. and b. Relationship between field measured and laser-estimated volume, c. magnitude of error in laser-derived estimates of merchantable stem volume of tree stem, and d. magnitude of error in estimates of merchantable stem length derived from laser-based reconstruction of individual tree stems. Trees scanned with standard pulse density are represented in red and those with high density in blue.

was also underestimated but the bias was smaller (mean -3.49%, std 6.37), and for 4 trees the bias was actually positive. For four trees the absolute volume error was below 2%. For all 14 trees with underestimation of merchantable volume greater than 4.5%, the stem reconstruction missed a portion of the upper part of the stem, and for 6 of those trees the missed portion was greater than 40% of the total merchantable stem length (Figure 3d). For several trees occluded portions of the stem were followed upwards the tree by portions that were successfully reconstructed (Figure 4). In those cases the volume of the missing part(s) could be retrieved by assuming a linear reduction in stem diameter between the upper diameter of the lower recovered stem section and the lower diameter of the upper section. Although this diameter interpolation improves the accuracy of the overall volume estimate it cannot be used to obtain reliable estimate of diameter that are needed for the development of volume equations.

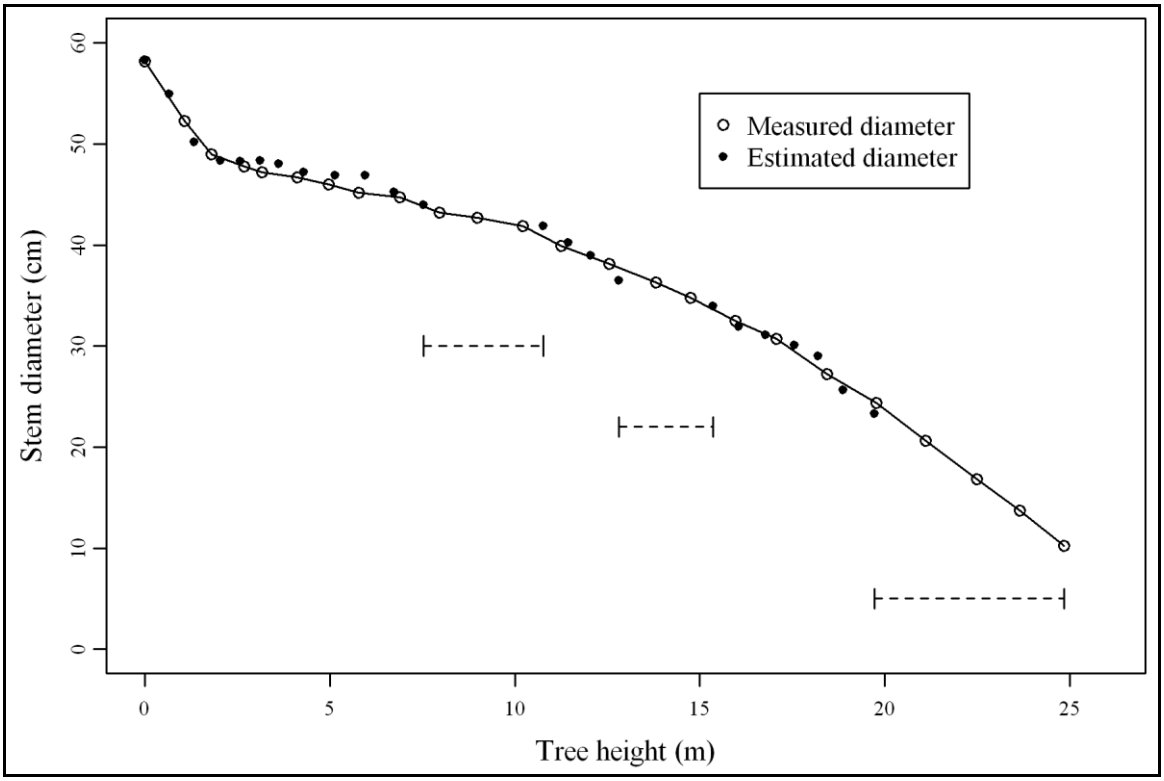

Figure 4. Laser-estimated and *in-situ* measured stem diameters at locations along the main stem of a ponderosa pine tree. Dash lines represent portions of the stem where reconstruction based on the laser point cloud failed. The continuous line linking measured diameters is for visualization purposes only.

These findings corroborate previous studies in which the upper parts of tree stems could not be reconstructed from terrestrial LiDAR data. When omnidirectional terrestrial LiDAR scanning is used to obtain a record of the plot or subplot vegetation structure, it is unlikely to expect that all trees in the scene will be reconstructed, except perhaps at leaf-off conditions. Where, however, only one tree is targeted for special purposes, including the development of localized volume equations, the operator of the laser instrument has a lot of flexibility in determining the location of the scanning stations, so as to maximize the chances for a successful, accurate, and precise reconstruction of the targeted stem. Experience from this study suggests that for a given pulse density at the tree level and tree species with predominantly horizontal branch structure, it is preferable to operate the scanner from a longer distance and higher density setting than from closely underneath with lower density. This is because in the former setup the pulse has to travel shorter distance through foliage and branches.

4. Conclusion

We evaluated the potential of terrestrial LiDAR as an alternative approach to destructive tree sampling for assessing the dimensionality of tree stems and for the development of volume equations. We managed to estimate the merchantable volume with acceptable amount of error (< 2%) and retrieve the entire length of the merchantable stem only for 2 of the 20 trees we considered, although even for those two trees, a few stems diameters had to be interpolated. At the present state of algorithmic development, an unobstructed section at or near the top of the merchantable tree stem is a prerequisite for obtaining an accurate estimate of stem volume using terrestrial LiDAR data.

5. References

Bienert, A., S. Scheller, E. Keane, F. Mohan, and C. Nugent. 2007. Tree detection and diameter estimations by analysis of forest terrestrial laserscanner point clouds. In *Proc. ISPRS Commission III Workshop*, 36(3):50-55.

Carey, V., L. Long, and R. Gentleman. 2010. RBGL: An interface to the BOOST graph library. R package version 1.24.0.URL http://CRAN.R-project.org/package=RBGL.

Chen, Q. 2007. Airborne lidar data processing and information extraction. *Photogrammetric Engineering and Remote Sensing*, 73(2): 109-112.

Cheng , Z.L., X.P. Zhang, and B.Q. Chen. 2007. Simple reconstruction of tree branches from a single range image. *J. of Computer Science and Technology*, 22(6):846-858.

Dijkstra, E.W. 1959. A note on two problems in connection with graphs. *Numerische Mathematik* 1:269–271.

Gentleman, E.W., W. Huber, and S. Falcon. 2010. graph: A package to handle graph data structures. R package version 1.26.0. URL http://CRAN.R-project.org/package=graph.

Hopkinson, C., L. Chasmer, C. Young-Pow, and P. Treitz. 2004. Assessing forest metrics with a ground-based scanning lidar. *Canadian J. Forest Research*, 34:573-583.

Lefsky, M., and M. McHale. 2008. Volume estimates of trees with complex architecture from terrestrial laser scanning. *J. of Applied Remote Sensing*, 2:023521, 19 pp.

Lumia, R., L. Shapiro, and O. Zuniga. 1983. A new connected components algorithm for virtual memory computers. *Computer Vision Graphics and Image Processing*, 22:287-300.

Max, T.A., and H.E. Burkhart, H.E. 1976. Segmented polynomial regression applied to taper equations. *Forest Science*. 22:283-289.

Pfeifer, N., B. Gorte, and D. Winterhalder. 2004. Automatic reconstruction of single trees from terrestrial laser scanner data. *Int. Archives of Photogrammetry and Remote Sensing*, Vol. XXXV, B5, pp. 114-119.

R Development Core Team. 2010. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Thies, M., N. Pfeifer, D. Winterhalder, and B.G.H. Gorte. 2004. Three-dimensional reconstruction of stems for assessment of taper, sweep, and lean based on laser scanning of standing trees. *Scandinavian J. Forest Research*, 19(6):571-581.

Verroust, A., and F. Lazarus. 2000. Extracting skeletal curves from 3D scattered data. *Visual Computer*, 16:15-25.

Xu, H., N. Gossett, and B. Chen. 2007. Knowledge and heuristic-based modeling of laser-scanned trees. *ACM Transactions on Graphics*, 26(4).